

Universidad San Andrés
M. Florencia Gabrielli

Maestría en Economía - 2022

Examen Domiciliario

Entrega: 26 de septiembre de 2022, por correo electrónico

Instrucciones:

- Trabajar en grupo de máximo dos personas para resolver los problemas numéricos está bien y se aconseja. Todos los miembros del grupo deben hacer los cálculos y entregarlos en forma individual, aunque aclarando que el trabajo fue hecho con otra persona. En un archivo de texto (word, látex, pdf) se debe escribir la respuesta completa.
- Cuando se le pide que informe los resultados, presente la respuesta en una tabla. Nada extravagante, pero no debe simplemente adjuntar una copia impresa del programa estadístico que utilizó. Un buen hábito es imitar las tablas de los papers en AER, por ejemplo. No hay nada peor que obtener resultados interesantes que son ilegibles para todos excepto pero usted. Debe adjuntar el código que utilizó para generar los resultados como un apéndice.
- El paper de Nevo [2000] es una buena guía que puede ser de utilidad para resolver la segunda parte de este examen.

GMM

Para el siguiente proceso generador de datos (dgp)

$$y = \beta x^2 + u$$

$$x = \pi z + v,$$

con $\beta = 1$, $\pi = 1$, y $z = 1$, donde (u, v) se distribuyen conjuntamente normal con medias 0 y varianzas 1 y correlación 0.8. Generar una muestra de tamaño 1000 y presentar en una tabla las estimaciones de β y el R^2 utilizando: OLS, NL2SLS, Two-Stage.

Estimación de Demanda: BLP

En el archivo adjunto se encuentra la base de datos que debe utilizar para hacer este examen. Los datos corresponden a una base armada con información sobre compras de medicamentos para el dolor de cabeza (i.e. aspirinas, etc.) en Barcelona. La data está relevada a nivel de tienda/semana para 3 marcas que vienen en 3 formatos (tamaños) diferentes y para una marca genérica (la marca 4) en dos tamaños. En total hay 11 combinaciones marca/tamaño que las llamamos simplemente “marca” en la base. Las otras variables en la muestra son:

- cantidad: cantidad de gente que va a la tienda cada semana.
- descuento: Si hay una promoción (oferta) en el producto esa semana.
- ventas.
- precio de venta del paquete
- costo del fabricante (proxy de precio mayorista)
- indicador de tienda.
- indicador de semana.

También tenemos una base con variables demográficas con la siguiente información:

1. ingreso*: ingreso del hogar para la persona *.
2. tienda: indicador de la tienda.

Un resumen de la base se muestra a continuación.

Cuadro 1: Estadísticas Descriptivas: Medicamentos para dolor de cabeza

Marca	Tamaño (tableta)	Market Share	Precio	Precio/50 tab	Precio mayorista unitario
Marca 1	25	8.90 %	3.52	6.91	2.26
Marca 1	50	11.10 %	4.99	4.99	3.76
Marca 1	100	7.60 %	7.14	3.57	5.93
Marca 2	25	7.30 %	3.02	6.29	2.02
Marca 2	50	5.10 %	5.19	5.19	3.70
Marca 2	100	2.20 %	8.23	4.12	6.25
Marca 3	25	2.50 %	2.66	5.35	1.90
Marca 3	50	2.00 %	3.81	3.81	2.67
Marca 3	100	4.90 %	4.05	4.05	3.73
Marca 4	50	6.20 %	3.57	3.57	2.48
Marca 4	100	4.20 %	3.29	3.29	1.51

1. Logit

Asuma la siguiente especificación para u_{ijt} , utilidad del hogar i por consumir la marca j en la semana t :

$$u_{ijt} = X_{jt}\beta + \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

donde X_{jt} son las características observadas de la marca j , ξ_{jt} es el parámetro que describe las características no observadas (por el econometrista) y ϵ_{ijt} es un término de error que es independiente e idénticamente distribuido (iid) entre hogares i y marcas j proveniente de una distribución logística.

Ayuda: haga esta parte en STATA, dado que será de utilidad para chequear el código de MATLAB luego. Asimismo, les aconsejo que nunca hagan limpieza de datos y manipulación de los mismos en MATLAB o cuestiones como regresiones con efectos fijos. En general se obtienen resultados poco precisos si uno no sabe usar todos los trucos que la gente usa para hacer estas cosas en MATLAB.

Realizar las siguientes estimaciones:

1. MCO con precio y promoción como las características de los productos.
2. MCO con precio y promoción como las características de los productos y dummies por marca.
3. MCO con precio y promoción como las características de los productos y dummies por marca-tienda (es decir la interacción entre marca y tienda).
4. Estimar los modelos 1, 2 y 3 utilizando el costo (proxy de precio mayorista) como instrumento.
5. Estimar los modelos 1, 2 y 3 utilizando el instrumento de Hausman (precio promedio de otras marcas en otros mercados).
6. Con la fórmula analítica para la elasticidad en un modelo logit, calcular las elasticidades-precio promedio de cada una de las marcas en el mercado, usando las estimaciones en 1,2 y 3. ¿Tienen sentido estos resultados? Argumente.

2. Modelo Logit con coeficientes aleatorios, i.e. BLP

Ahora considere el siguiente modelo de demanda con heterogeneidad individual en los coeficientes.

$$u_{ijt} = X_{jt}\beta + \beta_{ib}B_{jt}(\text{Producto de marca}) + \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

Los coeficientes aleatorios se determinan de la siguiente manera:

- Hay un coeficiente aleatorio en los productos de cada marca $\beta_{ib} = \sigma_B v_i$, donde v_i es una realización de la distribución normal estándar. (Notar que tenemos 4 marcas).
 - El coeficiente del precio depende del ingreso: $\alpha_i = \alpha + \sigma_I I_i$ (Ingreso).
- a) Estimar los parámetros del modelo β , α , σ_{ib} y σ_I utilizando el método BLP. Como instrumentos utilizar los costos y los precios del mismo producto en la misma semana en otras tiendas, no solamente el promedio sino una variable para cada precio en otra tienda en el mismo periodo en donde debe elegir otras 30 tiendas/mercados. Elegir como matriz de ponderaciones a la matriz óptima, $(Z'Z)^{-1}$. Si utiliza el truco de separar a los parámetros lineales de los no lineales para la estimación entonces debe tener dummies por marca

y promoción como características de los productos. Si no usa este truco, simplemente estime α , σ_{ib} y σ_I .

Utilice el comando *fminsearch* en MATLAB para minimizar la función objetivo de GMM, y utilice tanto 0 como algún otro valor como valor inicial (starting value). Simplemente reporte las estimaciones de los coeficientes y el valor optimizado de la función GMM. No hay que calcular los errores estándares para este ejercicio.

- b) Calcule las elasticidades, propias y cruzadas para la tienda 9 en la semana 10. ¿Cómo difieren estos números con respecto a las del modelo logit (simplemente asuma que los σ 's son 0 para ver esto). Argumente.
- c) Recupere los costos marginales para la tienda 9 en la semana 10 bajo el supuesto de que cada marca tiene un solo dueño. ¿Cómo difieren estos costos con respecto a los costos observados?

3. Análisis de fusión

Suponga que se fusionan las tres primeras marcas.

- a) Prediga los precios usando el modelo logit (sin coeficientes aleatorios) luego de la fusión pero solo para la tienda 9 durante la semana 10. Asegúrese que si no hay fusión, los precios que se obtienen no cambian!
- b) Escribir como haría para predecir el cambio en los precios usando el modelo de coeficientes aleatorios de BLP. No tiene que efectuar el cálculo, solamente decir cómo lo haría.

4. Guía para hacer BLP

El modelo de utilidad que se utiliza es

$$u_{ijt} = X_{jt}\beta + \beta_{ib}B_{jt} + \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

donde los coeficientes aleatorios son $\beta_{ib} = \sigma_B v_i$, con $v_i \sim N(0, 1)$ y $\alpha_i = \alpha + \sigma_I I_i$, donde I_i es el ingreso observado. Sustituyendo en la expresión para la función de utilidad y juntando los términos que representan la utilidad media del producto j , podemos reescribir el modelo de la siguiente manera

$$u_{ijt} = \delta_{jt} + \sigma_B v_i B_{jt} + \sigma_I I_i p_{jt} + \epsilon_{ijt}$$

donde la utilidad media derivada del producto j en la semana t es

$$\delta_{jt} = X_{jt}\beta + \alpha p_{jt} + \xi_{jt} \quad (1)$$

A los efectos de identificar y estimar los parámetros del modelo hacemos el siguiente supuesto $E[\xi_{jt}^0 Z_{jt}] = 0, \forall j, t$ para instrumentos Z_{jt} y para los verdaderos no observables ξ_{jt}^0 . Dado que no conocemos a los verdaderos no-observables, usaremos estimaciones $\xi_{jt}(\theta)$ donde $\xi_{jt} = \xi_{jt}^0$ evaluado en el verdadero valor del parámetro θ^0 . Los estimadores para los parámetros minimizarán la siguiente suma ponderada de cuadrados (con una matriz de ponderaciones óptimas),

$$\hat{\theta} = \arg \min_{\theta} \xi' Z (Z' Z)^{-1} Z' \xi \quad (2)$$

y para formar esta función objetivo necesitamos dos cosas: Z y ξ .

Para obtener ξ

- a) La ecuación (1) nos dice que $\xi_{jt} = \delta_{jt} - X_{jt}\beta - \alpha p_{jt}$. Notar que tenemos datos sobre X y p , pero aún necesitamos δ . [Berry, 1994] demuestra que existe un único vector δ que es solución de $s_{jt} = \tilde{s}(\delta_{jt}) \forall j, t$. BLP nos da la fórmula (contraction mapping) que permite resolver para δ ,

$$\delta_{jt}^{(m+1)} = \delta_{jt}^{(m)} + \log(s_{ij}) - \log(\tilde{s}(\delta_{jt}^{(m)}, \theta_2))$$

para esto necesitamos saber cómo calcular las participaciones de mercado que implica nuestro modelo, $\tilde{s}(\delta_{jt}, \theta_2)$.

- b) Las participación de mercado implicada por el modelo es la proporción de consumidores que eligen el bien j en el momento t (donde asumimos que ellos maximizan la utilidad)

$$\tilde{s}(\delta_{jt}, \theta_2) = \int \int \frac{\exp(\delta_{jt} + \sigma_B v_i B_{jt} + \sigma_I I_i p_{jt})}{1 + \sum_j \exp(\delta_{jt} + \sigma_B v_i B_{jt} + \sigma_I I_i p_{jt})} dF^v(v) dF^I(I).$$

El problema es que esta integral (múltiple) no tiene solución analítica, por lo tanto necesitamos simularla numéricamente. Esto es debemos simular ns “individuos” (para cada mercado t), cada uno caracterizado por el par (v_i, I_i) que se obtiene de las distribuciones apropiadas. Luego usamos el simulador

$$\hat{s}(\delta_{jt}) = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp(\delta_{jt} + \sigma_B v_i B_{jt} + \sigma_I I_i p_{jt})}{1 + \sum_j \exp(\delta_{jt} + \sigma_B v_i B_{jt} + \sigma_I I_i p_{jt})}$$

donde los ns elementos de la suma son las probabilidades de que el individuo i elija el producto j en el mercado t .

Instrumentos Z

Necesitamos instrumentos Z_{jt} que estén correlacionados con nuestros regresores (X_{jt}, B_{jt}, p_{jt}) pero no con ξ_{jt} . En este ejercicio utilizamos el precio mayorista (costo), el precio promedio en otros mercados y los precios en otros 30 mercados como instrumentos para el precio de la marca y, dejamos que los otros regresores (exógenos) sean sus mismos instrumentos.

Para obtener $\hat{\theta}$

Una vez que tenemos Z y ξ , la etapa de estimación consiste en resolver la ecuación (2) donde $\theta = (\beta, \sigma_B, \alpha, \sigma_I)$ son los coeficientes a estimar. Esta búsqueda no lineal se torna más compleja a medida que la dimensión del espacio de parámetros aumenta, de manera que un truco útil de realizar es “concentrar” en primer lugar a los parámetros lineales $\theta_1 := (\beta, \alpha)$ imponiendo las condiciones de primer orden de 2SLS (mínimos cuadrados en dos etapas) como se discute en el apéndice de Nevo [2000] y Berry, Levinsohn, and Pakes [1995].

Referencias

- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63(4), 841–890.
- BERRY, S. T. (1994): “Estimation of Discrete-Choice Models of Product Differentiation,” *Rand Journal of Economics*, 25(2), 242–262.
- NEVO, A. (2000): “A Practitioner’s Guide To Estimation of Random-Coefficients Logit Models of Demand,” *Journal of Economics and Management Strategy*, 9(4), 513–548.